

This is Youth Impact
Essay Collection



Is 'What works?' a question we should still be seeking to answer?

Nick Axford



Is 'What works?' a question we should still be seeking to answer?



Nick Axford PhD
Head of What Works
Dartington Social Research Unit

✉ nick.axford@dartington.org.uk
🐦 [@nick_axford](https://twitter.com/nick_axford)

Nick is a Senior Researcher and Head of What Works at the Dartington Social Research Unit. He leads a team that focuses on identifying effective interventions to improve child well-being through a combination of evidence reviews and evaluations (mostly randomised controlled trials).

Nick has been at the Unit since 1997. He is a member of the Early Intervention Foundation Evidence Panel and an Advisor to the Board of the European Society for Prevention Research, and was Co-Editor of the Journal of Children's Services from 2006 to 2015.

He tweets on evidence-based prevention and early intervention at @nick_axford.

An experiment was conducted last year in which 100 people were given pen pictures of 10 social interventions and asked to say whether each one would help, have no effect or cause harm.¹ The interventions ranged from boot camps and drug substitution programmes to school-based thinking skills lessons and mindfulness training to reduce stress. They were chosen because they were easy to explain and had been reviewed by the Campbell Collaboration², generating clear conclusions. Participants got four out of 10 correct on average – slightly better than a chimpanzee choosing at random. The authors, from 80,000 Hours, concluded:

"Sadly, it isn't actually possible for the public to know ahead of time whether a nice-sounding idea will actually help people or hurt them".

So for this and other reasons my answer to the question posed in the title is a resounding 'yes'.

- 1 For a longer description of the experiment, with an opportunity to take the test, see <http://www.vox.com/2015/8/13/9148123/quiz-which-programs-work>.
- 2 An international network that produces systematic reviews of the effects of interventions in areas such as crime and justice, education and social welfare: <http://www.campbellcollaboration.org>

First, as elegantly demonstrated by the experiment, some things that we think are good for children and young people are probably harmful. It is arrogant to think otherwise. The highest profile example in the recent past comes from the US. It seemed a good idea to take young people who had committed misdemeanors to an adult prison where they could confront the uncomfortable truth about life inside. Surely, it was reasoned, this would turn participants away from a life of crime: it would scare them straight.

Except that when results from trials of the programme (called 'Scared Straight') arrived, they showed something few had anticipated: not only did the intervention produce no benefits, it actually made matters worse.³ The control group kids offended more than the programme kids. (Only 15% of respondents in the experiment got this one right.) There are several other examples of interventions that increase crime or substance misuse, usually by bringing together groups of young people who

- 3 Petrosino, A., Turpin-Petrosino, C., Hollis-Peel, M. E. and Lavenberg, J. G. (2013) *Scared Straight and Other Juvenile Awareness Programmes for Preventing Juvenile Delinquency: A Systematic Review*. Campbell Systematic Reviews 2013:5.



“... some things that we think are good for children and young people are probably harmful. It is arrogant to think otherwise.”

display or are at risk of criminal behaviour.⁴

It begs the question: what else are we doing with good intentions for young people today that tomorrow we will find was woefully misplaced?

Second, not all interventions or services are equal.

At a glance, one parenting group looks much like another, and mentoring interventions by different names appear similar to the casual eye. Except that closer scrutiny reveals that not only are they often quite different but some are good at achieving their goals and others are not. The content, format and underlying theory can vary enormously. One might be rooted in helping participants to develop skills, whereas another might seek to alter their sense of self-efficacy or their perception of social norms. One may last 12 months and have a fixed set of topics to cover each week, while another may be much shorter and be led by what the participant wants to do or talk about. One might have been subjected to multiple trials, collectively showing consistent evidence of a positive impact, whereas another might have formative evidence only in the shape of feedback from focus groups indicating that users like it. The design, execution and write-up of studies that look at such things, and their appraisal against common standards, help interested parties to make distinctions that to the naked eye are high impossible.⁵

Third, interventions cost money and we need to know whether that money is being spent to best effect.⁶ The costs may range widely but even the cheapest intervention has an opportunity cost in terms of the time that users spend doing it when they could have been doing something else (potentially more worthwhile). Other costs typically cover staff or volunteer time, training, materials and ongoing support. Ultimately these are borne by the taxpayer, whether directly or indirectly, and of course users (or their parents) are often taxpayers themselves.

Most of us want public money to be spent well. All other things being equal, we would rather it is spent on hospitals not weapons, or on interventions that are effective and not those that are ineffective and will simply require further expenditure down the line. Moreover, faced with a choice between two interventions that are equally effective, most people would opt for the cheaper one, unless there are other good reasons not to.

“... we have a duty to promote young people’s well-being. Whether stated in terms of preventing or treating impairment to their health and development, or enabling them to flourish and lead fulfilling lives, it broadly amounts to the same thing.”

Fourth, we have a duty to promote young people’s well-being. Whether stated in terms of preventing or treating impairment to their health and development, or enabling them to flourish and lead fulfilling lives, it broadly amounts to the same thing. Any decent society should strive to meet its citizens’ needs, uphold their rights, protect their economic security, enhance their quality of life and enable their social inclusion.⁷ This includes protecting the vulnerable.

We have choices about how to do this, affected by

4 Rhule, D. M. (2005) Take care to do not harm: harmful interventions for youth problem behaviour. *Professional Psychology: Research and Practice* 36 (6), 618-625.

5 There is a danger of the field becoming flooded with standards, but two sets have become reasonably well established and are widely used. The first focuses on the quality of interventions, taking into account whether they are well-described, how rigorously they have been evaluated, the strength and consistency of their impact and how ready they are for dissemination via public service systems such as education and health (for the latest iteration see Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., Sandler, I. W. and Zafft, K. M. (2015) Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: next generation. *Prevention Science* 16 (7), 893-926). The second set focuses on the reporting of randomised controlled trials and requires that scientists are transparent about what they did and what they found (see Schulz, K. F., Altman, D. G., Moher, D. for the CONSORT Group (2010) CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332).

6 There are now online databases reporting on the costs and benefits of a range of interventions and types of intervention in the area of child and youth welfare. See www.wsipp.wa.gov and www.investinginchildren.eu.

7 There are arguably trade-offs between some of these, but arguably all are noble aims – see Axford, N. (2008) *Exploring Concepts of Child Well-being: Implications for Children’s Services*. Bristol: Policy Press.

considerations such as what is legal and moral, what people who use services want, and what is pragmatic (in terms of being affordable and deliverable). Scientific evidence of what is or is not effective - what works - is another important factor.

Fifth, the well-being of young people in the UK could and should be better, especially given that we are the fifth richest country in the world. Successive international league tables of child well-being have shown that our young people are worse off than their peers in many other developed countries on a range of objective and subjective indicators.

The latest report saw the UK placed 16th out of 29 countries overall, below the Czech Republic, Portugal and Slovenia.⁸ High rates of teenage pregnancy, alcohol abuse and young people not in education, employment or training were deemed particularly worrying. Similarly, a series of representative cross-sectional surveys in the UK covering the 30 years from 1974 to 2004 showed an increase in emotional and conduct problems among teenagers over that period (albeit with some plateauing and even possible reduction towards the end).⁹ Notwithstanding recent suggestions that today's teenagers seem reluctant to smoke, get drunk or have sex, perhaps unfairly earning them the moniker 'young fogeys'¹⁰, there is still scope for improvement and we need to learn how best to achieve this.

Sixth, when we expect young people or their families to invest time and effort participating in an intervention, it is only fair that we do so with a reasonable sense that it is worth their while. None of us would like to give up time to get a service only to be told later that it was a waste of effort. Worse still if we discover that as much was known before we took part.

If we are likely to be harmed, or if there is some doubt over the likely efficacy of what we are to be given, we deserve – and indeed are arguably entitled – to know, just as it is reasonable for us to know if there is evidence to support what we are being offered. We can then make an informed decision about whether or not to take part. It boils down to

8 UNICEF Office of Research (2013) *Child Well-being in Rich Countries: A Comparative Overview. Innocenti Report Card 11*. Florence: UNICEF Office of Research.

9 Nuffield Foundation (2009) *Time Trends in Adolescent Well-being: Update 2009*. London: Nuffield Foundation.

10 Margolis, E. (2016) The new young fogeys. *New Statesman*, 8th February. <http://www.newstatesman.com/politics/uk/2016/02/new-young-fogeys>

“When we ask ‘What works?’ we are using short and simple words to convey rich and complex meaning. It is not like flicking a light switch and asking if the bulb works.”

the simple matter of respect.¹¹

Seventh, in many areas of services to promote young people's well-being we know relatively little about what works, and the more we explore the more uncharted territory comes into view. The frequent conclusion of authors who systematically review the 'what works' evidence on any given subject is 'More research is needed'.

This is not, usually, the thinly veiled plea for extra research funding that it might seem. Rather, it reflects the reality that when we focus on a specific type of intervention for a discrete population and problem, and if we narrow our criteria to include only certain evaluation designs and then critically appraise the studies, it is often difficult to draw firm conclusions about what is effective. Even if some general messages emerge, there will be unanswered questions about why something is effective, for whom and in what circumstances. Related to this, we are constantly faced with new problems that need investigating, new interventions and even new types of intervention, including some that extend the focus beyond the individual young person and their family and seek to change the wider ecosystem.

So yes, we still need to answer the question ‘What works?’ (indeed, some would say we have barely started to ask it). But the question and associated assumptions are not unproblematic.

When we ask ‘What works?’ we are using short and simple words to convey rich and complex meaning. It is not like flicking a light switch and asking if the bulb works. We mean at least what works for whom, when and where. In other words, does this course of action work for everyone or just some people, and does

11 Farrington, D. P. (2003) A short history of randomized experiments in criminology: a meager feast. *Evaluation Review* 27, 218-227.



it work when tried in this context as well as in that context?

It is widely accepted that the average effect of interventions masks considerable variety: some recipients make leaps and strides, while others emerge largely untouched or even go downhill. We need to tease out such differential effects and find out who benefits most and why. In some cases, for instance, young people with the most serious problems at the outset make the greatest progress, perhaps because they have the furthest to travel (it is harder for those with fewer problems to get much better). Similarly, the same type of intervention may be effective in some settings but not others, reflecting variability in cultural norms, implementation infrastructure, the nature of 'services as usual' and the wider socio-political climate.

“When we ask ‘What works?’ we are also investigating the mechanisms that produce the effect... ‘What works’ includes what makes something work. ”

While it is true that ‘what works there’ doesn’t automatically translate into ‘what works here’¹², it does not follow that ‘what works here’ must only ever be established by testing something in situ. We can extrapolate, particularly if there is breadth and depth in other studies and no clear evidence of context moderating impact. Put another way, if multiple studies across different settings show similar effects, we can be more confident.

When we ask ‘What works?’ we are also investigating the mechanisms that produce the effect. Sometimes people refer to this as inquiring about *how* or *why* something works. It is arguably part of the same thing but at a different level of detail. We know a car works insofar as it moves when you press the accelerator, but lift the bonnet and we find it is a bit more complicated: motion comes from igniting a mix of fuel and air in a combustion chamber to create an explosion that pushes pistons which in turn drive the crankshaft. ‘What works’ includes what makes something work. It is common for intervention developers to design their

product around a logic model or theory of change.¹³ This is the skeleton of the intervention, a hypothesis about how certain activities will produce the desired outcomes.

The best logic models spell out clear chains of effect that can be tested. Perhaps our prediction will turn out to be wrong: an intervention may achieve its goals but via a different route, or fail to achieve its goal because a link in the chain is faulty. This is very useful information, as it can be applied to developing new interventions or tweaking the one in question. Without opening the black box we are little the wiser, yet to date relatively few studies in our field have attempted this.

When we ask ‘What works?’ we don’t spell out what we mean by ‘what’, which can be myriad. For some time now the term has arguably been associated with programmes – discrete and organised packages of activities spelled out in a manual and accompanied by training and technical assistance.

But intervention can take different forms – policies such as minimum unit pricing for alcohol, practices such as cognitive behavioural therapy or processes such as screening. Deliberately adjusting aspects of the environment to shape people’s behaviour choices is another approach, sometime referred to as ‘nudge’¹⁴, as is the identification and dissemination of common units of behavioural influence or ‘kernels’¹⁵ that practitioners may improvise with depending on the needs of the user. These latter approaches are relatively new, and there is much to learn about their feasibility, let alone their impact.

When we ask ‘What works?’ we are asking a question, not making a statement about the evaluation method, yet the two are often confused. The task of exploring ‘what works’ does not necessarily entail large long-term studies with comparison groups in which all aspects of the method are spelled out up front in a ‘protocol’ from which deviation is a sin. Nor does it presume that the only worthwhile research activity is crunching data from multiple such studies (known as ‘meta-analyses’).

12 Cartwright, N. and Hardie, J. (2012) *Evidence-based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press.

13 Funnell, S. C. and Rogers, P. J. (2011) *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco, Jossey-Bass.

14 Thaler, R. H. and Sunstein, C. R. (2008) *Nudge: Improving Decisions about Health, Wealth and Happiness*. New Haven: Yale University Press.

15 Embry, D. D. and Biglan, A. (2008) Evidence-based kernels: fundamental units of behavioral influence. *Clinical Child and Family Psychology Review* 11 (3), 75-113.

“We can afford to be more catholic in our approach, recognising that seeking to enhance young people’s well-being is a complex business.”

Whisper it quietly, but finding out what works is also not something that can only be done with randomised controlled trials, or RCTs, important and valuable though they are. We can afford to be more catholic in our approach, recognising that seeking to enhance young people’s well-being is a complex business.

Indeed, while ‘proving’ impact has its place, there is a growing emphasis on ‘improving’ – making small adjustments to what is done, testing what effect this has, harvesting the lessons, and repeating this in a cyclical process.¹⁶ Other methods besides trials – focus groups, interviews, observation – come into their own here. This is not just about impact on outcomes, indeed much attention should be paid to traction in terms of deliverability and ‘pull’ – do practitioners deliver the intervention, what do they adapt and why, do they like it, do users use it and what would make them want to use it more? The aim is to move on from the failed approach of waiting to find out if an intervention is effective before thinking about scalability. Terms like ‘rapid cycle testing’ and ‘non-sequential evaluation’ are increasingly used to capture these ideas.

When we ask ‘What works?’ we are also in danger of sounding incredibly naïve. As indicated already, this is partly because of the complexity involved. Psychosocial interventions themselves are invariably complex, constituted as they are of multiple moving components. The environments into which they are thrust are also messy, comprising multiple interacting systems ranging from the family to the prevailing political and ideological climate. Then there are individuals themselves, complicated bundles of

genetic make-up and experience.

Given this, is it really possible to isolate what works – to isolate active ingredients, to attribute effect to cause? Some commentators argue that this is why RCTs are of limited value¹⁷, but the counter-argument would be that it is precisely why they are important: they control for observed and unobserved factors that bear on the outcome in question.¹⁸ It is the complexity involved that renders them an essential tool.

“The aim is to move on from the failed approach of waiting to find out if an intervention is effective before thinking about scalability.”

Lastly, when we ask ‘What works?’ the question is inherently loaded owing to the politically and morally contested nature of the means and ends involved. We might find, for example, that locking children in a prison cell for 23 hours a day stops them from committing crime (it is hard to imagine that it wouldn’t, at least while they are there). But few people of a liberal progressive mindset would consider this justifiable. Equally, we may discover that a drug misuse prevention programme makes young people more risk-averse, but is this desirable if, as we know, some risk-taking behaviour is functional, either in the short-term (a good way to bond and make friends) or the long-term (people with highly successful careers often display such behaviours)?

In some ways there is relatively little that can be done about these issues, other than to be alert to their danger, but it should also be acknowledged that some ends are arguably universal goods that deserve to be pursued (they are less contested than some might claim).¹⁹

16 For a discussion of some of these issues see Dartington Social Research Unit (2015) *Evidence is Confidence: How to Create a Richer Evidential Tapestry*. Realising Ambition Programme Insight 2. http://dartington.org.uk/inc/uploads/Realising_Ambition_Programme_Insight_Evidence_is_Confidence.pdf. The website of the Institute for Healthcare Improvement has some useful overviews and tools on the science of intervention improvement, including rapid cycle testing and the PDSA (Plan-Do-Study-Act) cycle: <http://www.ihl.org/resources/Pages/HowtoImprove/ScienceofImprovementHowtoImprove.aspx>.

17 For example, see Pawson, R. (2013) *The Science of Evaluation: A Realist Manifesto*. London: Sage Publications.

18 Bonell, C., Fletcher, A., Morton, M., Lorenc, T. and Moore, L. (2012) Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Social Science & Medicine* 75 (12), 2299-2306.

19 Doyal, L. and Gough, I. (1991) *A Theory of Human Need*. New York: Guilford.



The question posed in the title implies that asking 'What works?' might have had its day. That perhaps we know enough, or that there is no scope to improve young people's well-being, or that the question can't be answered. That it is time to stop and move on. I beg to differ.

Asking 'What works?' comes from a position of curiosity and humility. We acknowledge the poverty of our knowledge and don't presume to have the answers, but we want to find out, to learn.

Why? Because we are humane: we want to minimise human suffering, increase human flourishing, and help create a better and more just society for all.

This is Youth Impact **Essay Collection**

This essay forms part of a larger collection, published by the Centre for Youth Impact. The collection was published to take stock of the evidence and impact debate in the youth sector, share different perspectives, and reflect on where the conversation could take us.

About the Centre for Youth Impact

The Centre for Youth Impact is a community of organisations working together to increase thinking and practice around impact measurement in work with young people.

We offer:

- A 'route in' to information, support and debate in relation to evidence and impact
- Events where you can collaborate with others and learn more about the impact agenda
- Resources to help improve impact measurement practice
- An inclusive platform promoting debate and discussion to advance both thinking and practice

Learn more or take part

-  www.youthimpact.uk
-  hello@youthimpact.uk
-  Follow @YouthImpactUK